

## RESEARCH

## Open Access



# Methicillin-resistant *Staphylococcus aureus* emerged long before the introduction of methicillin into clinical practice

Catriona P. Harkins<sup>1,2</sup>, Bruno Pichon<sup>3</sup>, Michel Doumith<sup>3</sup>, Julian Parkhill<sup>4</sup>, Henrik Westh<sup>5,6</sup>, Alexander Tomasz<sup>7</sup>, Herminia de Lencastre<sup>7,8</sup>, Stephen D. Bentley<sup>4</sup>, Angela M. Kearns<sup>3</sup> and Matthew T. G. Holden<sup>1,4\*</sup>

## Abstract

**Background:** The spread of drug-resistant bacterial pathogens poses a major threat to global health. It is widely recognised that the widespread use of antibiotics has generated selective pressures that have driven the emergence of resistant strains. Methicillin-resistant *Staphylococcus aureus* (MRSA) was first observed in 1960, less than one year after the introduction of this second generation beta-lactam antibiotic into clinical practice. Epidemiological evidence has always suggested that resistance arose around this period, when the *mecA* gene encoding methicillin resistance carried on an SCCmec element, was horizontally transferred to an intrinsically sensitive strain of *S. aureus*.

**Results:** Whole genome sequencing a collection of the first MRSA isolates allows us to reconstruct the evolutionary history of the archetypal MRSA. We apply Bayesian phylogenetic reconstruction to infer the time point at which this early MRSA lineage arose and when SCCmec was acquired. MRSA emerged in the mid-1940s, following the acquisition of an ancestral type I SCCmec element, some 14 years before the first therapeutic use of methicillin.

**Conclusions:** Methicillin use was not the original driving factor in the evolution of MRSA as previously thought. Rather it was the widespread use of first generation beta-lactams such as penicillin in the years prior to the introduction of methicillin, which selected for *S. aureus* strains carrying the *mecA* determinant. Crucially this highlights how new drugs, introduced to circumvent known resistance mechanisms, can be rendered ineffective by unrecognised adaptations in the bacterial population due to the historic selective landscape created by the widespread use of other antibiotics.

**Keywords:** *Staphylococcus aureus*, MRSA, Antibiotic resistance

## Background

Methicillin-resistant *Staphylococcus aureus* (MRSA) has been identified as one of the major risk pathogens associated with the development of antimicrobial resistance (AMR). The emergence of AMR in *S. aureus* is well documented and the species has proven particularly adept at evolving resistance in the face of new antibiotic challenges. The introduction of penicillin in the 1940s heralded a revolution in the treatment of infectious diseases. However, at the same time as its use was

becoming more widespread following advances in the scaling up of production, evidence of penicillin resistance in *S. aureus* was already being uncovered [1].

Methicillin (Celbenin), a semi-synthetic  $\beta$ -lactam, was introduced in the UK in 1959 to circumvent growing penicillin resistance in *S. aureus*, associated with the acquisition of a  $\beta$ -lactamase enzyme, *blaZ* [2]. As a second-generation  $\beta$ -lactam antibiotic, methicillin was insensitive to breakdown by BlaZ. Following the introduction of methicillin into clinical practice in the UK, the Staphylococcal Reference Laboratory in Colindale (London, England) screened *S. aureus* isolates for evidence of resistance to this antibiotic [3]. More than 5000 *S. aureus* strains were assessed in the period

\* Correspondence: [mtgh@st-andrews.ac.uk](mailto:mtgh@st-andrews.ac.uk)

<sup>1</sup>School of Medicine, University of St Andrews, St Andrews KY16 9TF, UK

<sup>4</sup>The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK

Full list of author information is available at the end of the article



between October 1959 and November 1960, and in October 1960 three isolates showing increased minimum inhibitory concentrations (MICs) to the new drug, methicillin, were identified. The isolates originated from the same hospital and shared a common phage type and resistance profile (penicillin, streptomycin, and tetracycline), suggesting that they were related. In the description of these isolates it was noted that methicillin had been used only once previously at this hospital, and that none of the individuals from whom MRSA was isolated had been exposed to the drug. Within 2 years MRSA was being detected elsewhere in Europe, with invasive infections being identified in Denmark [4]. These MRSA isolates from the UK and Denmark in the early 1960s constitute the very first epidemic MRSA clone.

The genetic basis of methicillin resistance in *S. aureus* is associated with carriage of a mobile cassette of genes known as the staphylococcal cassette chromosome *mec* (SCC*mec*) [5]. Within this cassette is the *mecA* gene that is responsible for resistance to  $\beta$ -lactams including methicillin. The product of *mecA* is the peptidoglycan synthesis enzyme penicillin binding protein (PBP) 2a involved in cross-linking of peptidoglycan in the bacterial cell wall [6, 7]. PBP2a has a lower binding affinity for  $\beta$ -lactam antibiotics than the native PBP proteins encoded in the core genome of *S. aureus*. The subsequent combination of reduced penicillin-binding affinity and increased production of PBP2a accounts for the observed resistance to  $\beta$ -lactam antibiotics.

Genetic analyses of the first MRSA by multi-locus sequence typing (MLST) demonstrated that they were sequence type (ST) 250, a lineage belonging to clonal complex (CC) 8 and carried the type I SCC*mec* element [8, 9]. After emerging in the UK, this first epidemic MRSA clone (ST250-MRSA-I) spread across Europe during the 1960s and 70s, but by the late 1980s had become less prevalent and is now rarely reported [9–11]. The single locus variant and close relative of ST250-MRSA-I, ST247-MRSA-I was first detected in Denmark in 1964 [8] and has been more successful, spreading globally and persisting as a source of outbreaks in Europe into the late 1990s [10, 11], but this too has been superseded by more successful contemporary clones [10]. Five decades on since the appearance of the first MRSA, multiple MRSA lineages have emerged which have acquired different variants of SCC*mec* elements.

Epidemiological evidence has always suggested that MRSA arose as a consequence of the introduction of methicillin into clinical practice. Here we have used whole genome sequencing of a collection of 209 of the earliest MRSA isolates recovered in Europe between 1960 and 1989 to reconstruct the evolutionary history of methicillin resistance. Using Bayesian phylogenetic reconstruction we have identified the likely time point at

which this early lineage arose and also predicted the time around which SCC*mec* was acquired.

## Results

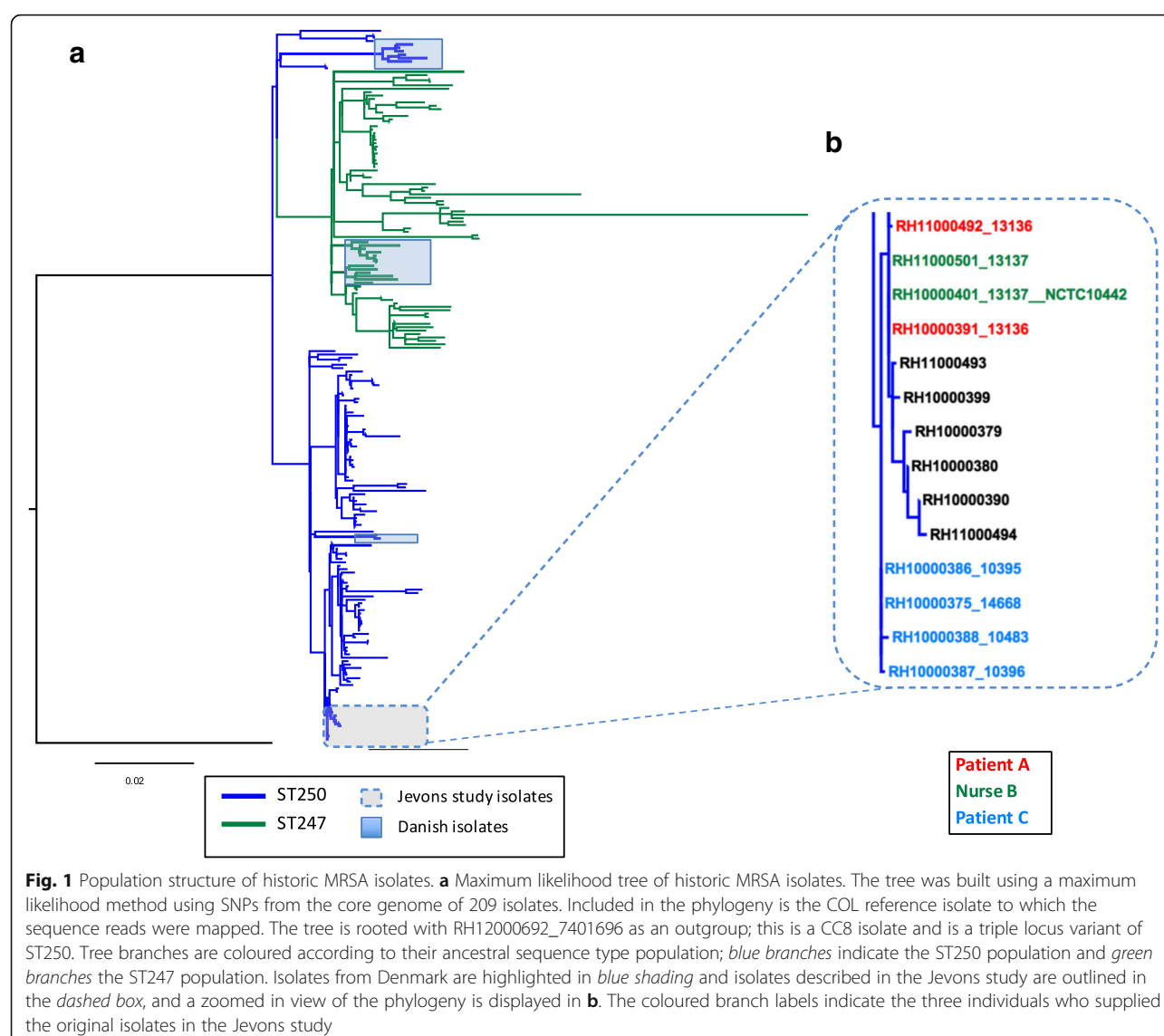
### Early MRSA belong to a diverse clone

Preserved in the culture collection of the Staphylococcal Reference Laboratory at Public Health England are representatives of the very first MRSA identified. These original isolates have been preserved as freeze-dried cultures, and have not been repeatedly passaged over the years. One hundred and eighty eight isolates that represented the earliest MRSA were recovered from the ampoules and their genomes sequenced (Additional file 1: Table S1). All the isolates belonged to CC8 MRSA and were originally isolated between 1960 and the late 1970s, and included eight isolates from the original study describing MRSA in 1961 [3]. In addition, 21 CC8 MRSA isolated between 1964 and 1989 in Denmark [8, 11] were sequenced, as representatives of the earliest MRSA detected elsewhere in Europe. We also included early methicillin-sensitive isolates of ST250 or ST247 ( $n = 11$ ); however, only a limited number of these were found in the reference laboratory collection.

Analysis of the MLST of the isolates identified two main groups, ST250 ( $n = 126$ ) and a single locus variant (SLV), ST247 ( $n = 78$ ), plus two novel SLVs of ST247 ( $n = 4$ ) (Additional file 1: Table S1). A supplementary isolate from the Public Health England collection was included to provide an outgroup for the analysis; RH12000692\_7401696 is an MRSA which was collected in 1967 and is a triple locus variant of ST250 (Additional file 1: Table S1).

The *S. aureus* isolate COL, a representative member of this early MRSA lineage first identified in the 1960s [12], had previously been fully sequenced, and the chromosome was used as a reference for mapping. Following exclusion of mobile genetic elements (MGEs) and predicted recombination events in the collection, a total of 4220 SNPs were identified and used to construct a phylogeny (Fig. 1a). The population framework revealed a diverse population structure containing several distinct clades. The mapping of the ST information on to the phylogeny reveals that the ST250 population is basal to the ST247, suggesting that ST247 emerged from ST250, which is consistent with the epidemiological evidence, and supports the hypothesis that this pandemic multidrug-resistant MRSA clone emerged out of the ancestral MRSA genotype [8, 9].

Highlighted in the expanded view (Fig. 1b) are the isolates from the Jevons study, derived from three individuals at the same hospital in the South London area between July and November 1960 [3]. The isolation source and resistance profiles of these isolates are shown in Additional file 2: Table S2. These isolates are



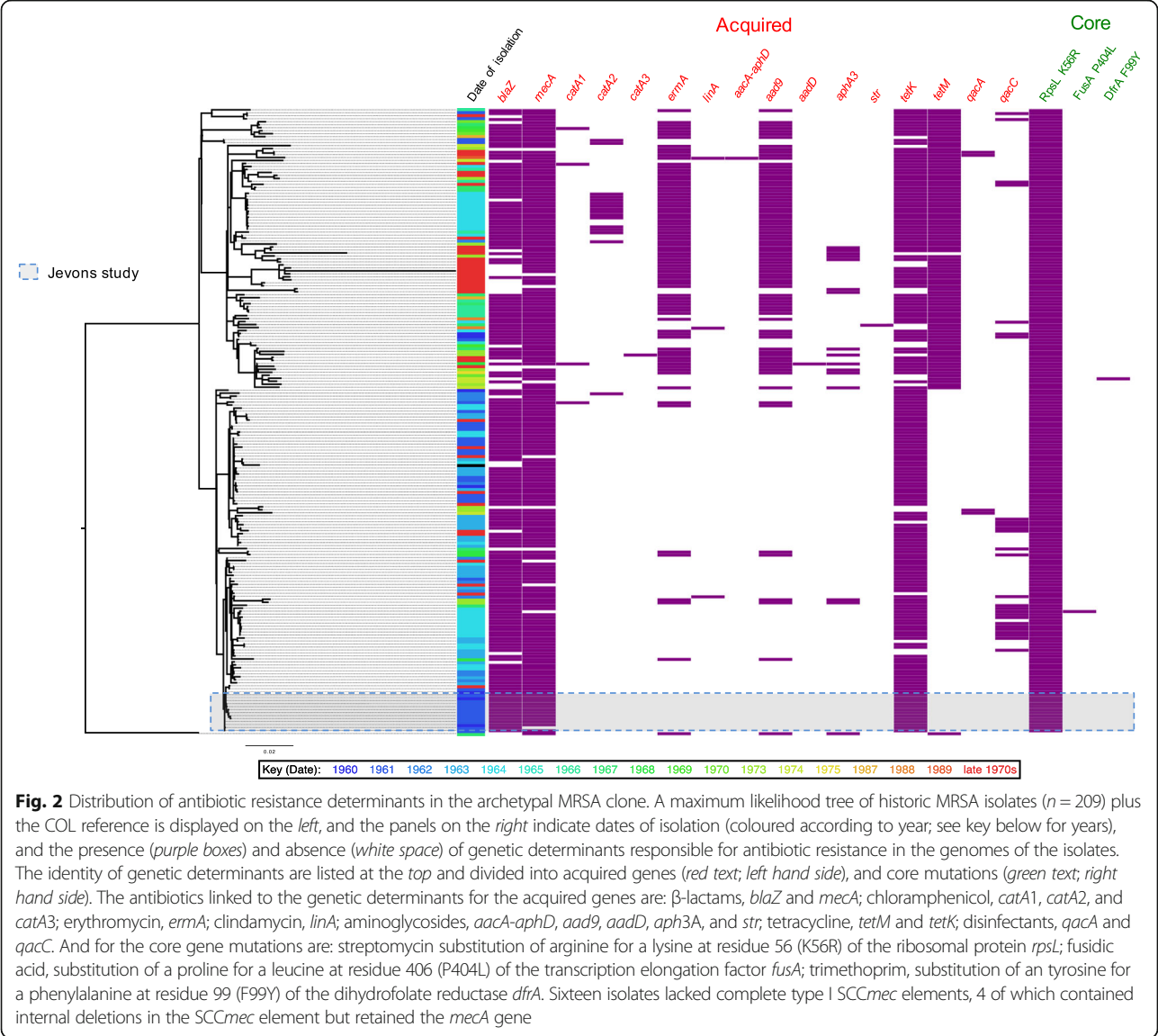
genetically very closely related, differing by seven SNPs only. Present within this cluster are additional isolates from the Public Health England collection originating between 1960 and 1961. Full epidemiological data are not available for these, but two of these isolates were identified in the same region as the hospital where the original Jevons study isolates originated. The genetic distance between isolates and their phylogenetic relationships suggests there was transmission within the hospital between patients A and C and nurse B, and that they were also transmitted beyond the hospital as part of a local outbreak.

Although all of the Jevons isolates are confined to a single clade, other isolates from the early 1960s are distributed throughout the entire phylogeny (Fig. 2). This suggests that the earliest MRSA circulating in the UK were not from a single recently emerged clone, but

were representatives of an established population. In addition to the UK isolates, there were 21 from Denmark, which represent the earliest MRSA detected outside the UK. These derive from 1964 onwards, and include the youngest isolates within the collection from the late 1980s. The Danish isolates are found in three clusters distributed throughout the phylogeny (Fig. 1a), suggesting that, like the early UK MRSA, they originated from an established and diverse population.

#### Genetic basis of resistance to methicillin and other antibiotics in the archetypal MRSA population

Previous studies have shown that the archaic MRSA clone carried a type I *SCCmec* element, which was the first type of this MGE family to be classified [5, 13]. Notably, the description of the type I element was based upon the *SCCmec* derived from *S. aureus* strain



NCTC10442 identified in the 1960 Jevons study (Fig. 1b; Additional file 2: Table S2) [13]. The type I element carries *mecA* as its only resistance gene in combination with a truncated gene encoding the MecRI regulatory proteins (together known as a class B *mec* gene complex) with type 1 chromosomal recombinases (*ccrA1* and *ccrB1*). The original description of SCCmec type I identified the presence of a frameshift mutation in *ccrB1* which disrupts the translation of this site-specific recombinase [13]; the mutation occurs after codon 321 and is caused by a single base deletion in a poly-A hexamer resulting in a pentamer sequence. In the collection, 193 of the isolates contained intact SCCmec elements carrying the *mecA* gene (Fig. 2). Of these, 192 were SCCmec type I elements, all of which contained the pentamer sequence and the same frameshift mutation in *ccrB1* as

the NCTC10442 reference. The only non-type I element identified in the collection was in the outgroup isolate RH12000692\_7401696, which contained a type IVh SCCmec element. The remaining 16 isolates lacking complete SCCmec elements were distributed throughout the phylogeny, suggesting that these represent methicillin-sensitive *S. aureus* (MSSA) arising from the loss of the type I SCCmec element, rather than forming an ancestral MSSA population.

In addition to methicillin resistance, the first MRSA described were also resistant to penicillin, streptomycin and tetracycline [3]. Analysis of the genomes of these isolates identified *blaZ* and *tetK* genes conferring resistance to penicillin and tetracycline, respectively, but failed to identify the *str*, *aadE* or *aad9* genes associated with streptomycin resistance in *S. aureus*. In the absence of



an acquired resistance gene, the core genome was examined for mutations potentially responsible for resistance to streptomycin. In *Mycobacterium tuberculosis*, mutations in the ribosomal protein RpsL were shown to confer streptomycin resistance, including the substitution of an arginine for a lysine residue at residue 43 [14]. Alignment of the *M. tuberculosis* and *S. aureus* sequences revealed that RpsL in the Jevons isolates contained an arginine in the equivalent position, residue 56. Comparison with RpsL sequences in the public sequence databases showed that in *S. aureus* the frequent amino acid residue at position 56 was lysine. Examining the whole collection, all but one of the sequenced isolates contained the arginine residue at position 56, the exception being the outgroup isolate RH12000692\_7401696 (Fig. 2). This demonstrates that the non-synonymous substitution resulting in an arginine for a lysine residue at residue 56 (K56R) occurred most likely very early during emergence of the archetypal MRSA population.

*In silico* analysis of the resistomes of the isolates revealed genetic resistance determinants to numerous other antibiotics, including penicillin (*blaZ*), erythromycin (*ermA* and *linA*), kanamycin (*aadD*), gentamicin and kanamycin (*aacA-aphD*), spectinomycin and streptomycin (*aad9*), and chloramphenicol (*catA1*, *catA2* and *catA3*), fusidic acid (*fusA* P404L) and trimethoprim (*dhfrA* F99Y), as well as genes associated with decreased susceptibility to disinfectants (*qacA* and *qacC*). The frequency and widespread dispersal of these determinants reveal the strong selective pressure exerted by antibiotics on the archetypal MRSA clone over an extensive period. Examining their distribution in the context of the phylogeny shows that some of these traits have been co-acquired (Fig. 2), such as *ermA* and *aad9*, which are carried on Tn554, and that these acquisition events can be mapped on to the phylogeny [15].

### Evolution and emergence of methicillin resistance

To determine if the methicillin resistance emerged once or multiple times in the archetypal MRSA population, we examined the variation within the *SCCmec* type I elements. In total, 194 variant sites were identified in 192 elements present in the collection. Analysis of the distribution of the variation within the elements suggested that some could be attributed to homologous recombination. Two regions contained the majority of the variation: 124 SNP sites were identified in the gene encoding the LPxTG surface protein *pls*, and 31 SNP sites within a 549-bp intergenic region between a hypothetical protein (SACOL0030) and a glycerophosphoryl diester phosphodiesterase (SACOL0031). Excluding these predicted recombination regions, 39 core variants sites across 28.6 kb distinguished the 192 elements, with half of the isolates ( $n = 96$ ) carrying an identical element.

The maximum SNP distance distinguishing any two elements was eight SNPs, and phylogenetic analysis revealed that the elements present in the historic MRSA clone were closely related (Additional file 3: Figure S1) and shared a common evolutionary origin.

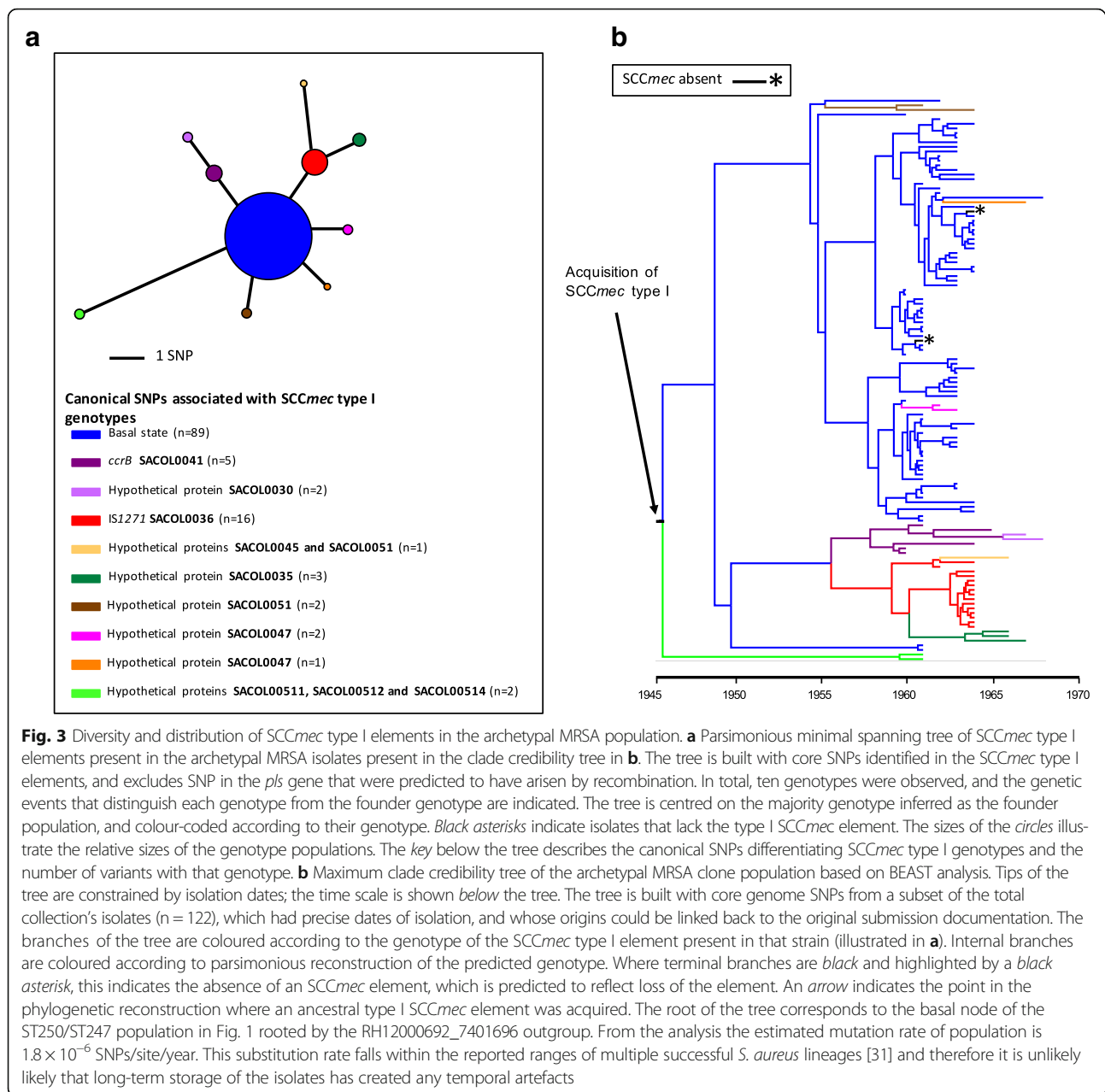
Our analysis of the evolutionary events surrounding the emergence of methicillin resistance in the archetypal MRSA lineage focused on a subset of 122 isolates that had precise dates and places of origin which could be linked to original submission records (Additional file 2: Table S1). This enabled us to generate a robust Bayesian phylogeny and temporal calibration. Examining the distribution of the type I *SCCmec* variants (Fig. 3a) within the context of a core genome phylogeny generated with BEAST (Fig. 3b) revealed congruence between the phylogenetic relationships of the two. All of the canonical SNPs associated with the *SCCmec* genotypes could be singularly mapped onto nodes of the core phylogeny, suggesting that the variation observed in the *SCCmec* elements had occurred during expansion of the ST250 and ST247 populations. On the basis of this, we propose that a type I *SCCmec* element was acquired once in a single primordial development of methicillin resistance (Fig. 3b) that could be dated back to the emergence of this clone.

In our Bayesian phylogenetic analysis of the core genome SNPs we utilized a range of population and clock model combinations. The combination of an exponential population and relaxed-log normal clock model was found to be the best fit to our data based on Bayes factors using the harmonic mean estimator. This indicated the time to most recent common ancestor (TMRCA) of the ST250/ST247 population was 1946 (95% highest posterior density (HPD) 1938–1952) (Additional file 3: Figure S2), and therefore the time of acquisition of *SCCmec* was likely around, or before, this date. Notably, the TMRCA of the type I *SCCmec* elements in these isolates based on a linear regression of a core SNP phylogeny was predicted to be early 1941 (Additional file 3: Figure S3).

To ensure that the Bayesian result was not an artefact of the clock or population models used in the analysis, we calculated the TMRCA for a range of model combinations and found that our chosen model exhibited a predicted TMRCA that was encompassed by the 95% HPDs of all other model combinations (Fig. 4).

### Discussion

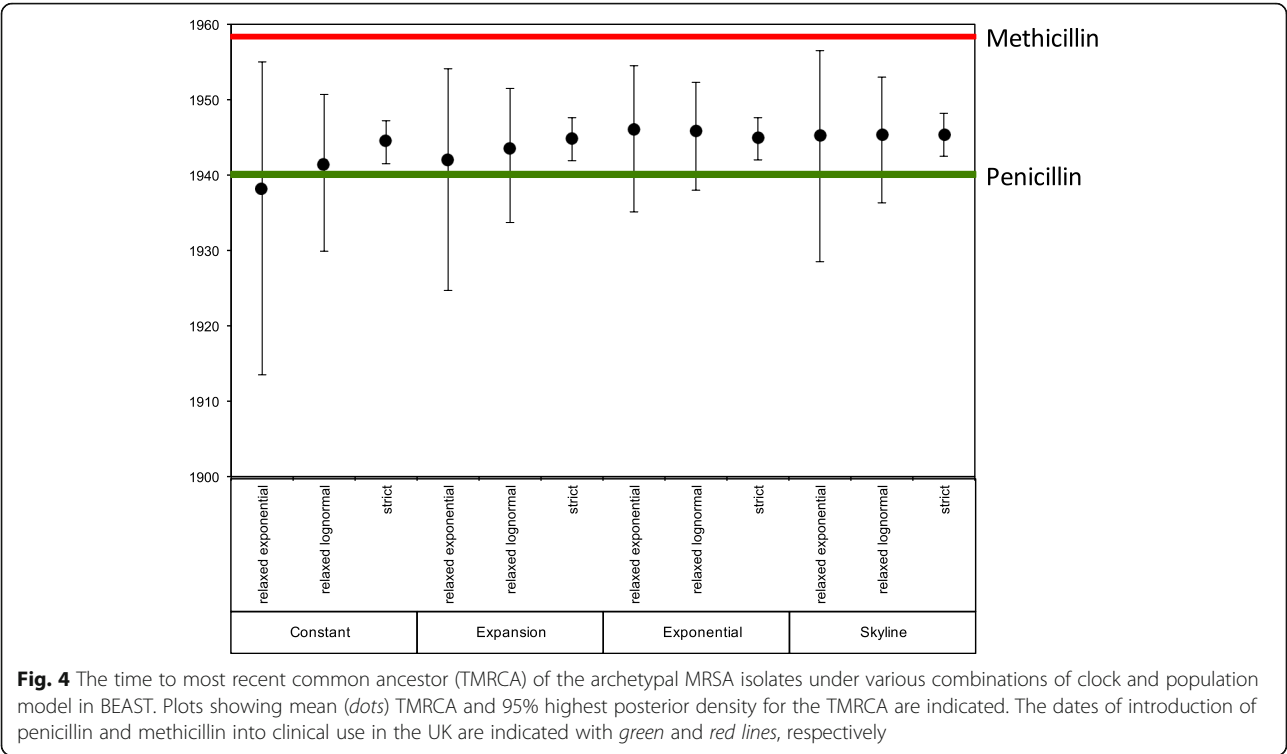
This historic collection provides unique insights into of the evolution of the first MRSA lineage. Preserved for decades in their original freeze-dried state, this large collection of strains representing the earliest MRSA clone has allowed us to reconstruct the evolutionary events leading to the emergence of MRSA. Using whole



genome sequencing we have been able to shed light on the time when SCCmec first entered into *S. aureus*, and also to estimate how many times this is likely to have happened in the archaic MRSA population.

The origins of SCCmec almost certainly lie in coagulase-negative staphylococci (CoNS) [16]. *S. aureus* belonging to the ST250 background appear to have been the first recipient in the transfer from CoNS, but whether the element entered the ST250 population on multiple occasions, or as a single isolated event with subsequent propagation through the population, has never been definitively resolved. A single entry of *mecA* followed by its evolution within the recipient

background has been suggested [17]. In order to clarify this we examined the variation present within the SCCmec elements in isolates throughout the population. The variation seen within SCCmec is predominantly in the *pls* gene, which has been described before [18]. Functionality of this 230 kDa cell wall-anchored (CWA) protein remains unclear, but its expression has been shown to reduce adhesion to host proteins as well as decrease invasiveness [19]. This LPxTG surface protein has a highly repetitive D/S-rich structure, making it a target for homologous recombination. As noted in other lineages the CWA proteins are subject to diversifying selection and exhibit diversity between and within *S.*



**Fig. 4** The time to most recent common ancestor (TMRCAs) of the archetypal MRSA isolates under various combinations of clock and population model in BEAST. Plots showing mean (dots) TMRCAs and 95% highest posterior density for the TMRCAs are indicated. The dates of introduction of penicillin and methicillin into clinical use in the UK are indicated with green and red lines, respectively

*aureus* lineages [20, 21]. Removal of this variation reveals that the evolutionary history of the SCCmec elements was congruent with that of the strains carrying them, which points towards a single acquisition, rather than multiple or recurrent horizontal transmissions. Supporting this hypothesis is the observation of a mutation in *ccrB1* gene of the SCCmec type I element. The recombinase genes are required for both integration and excision from the chromosome. Specifically, CcrB is required for excision and the mutation present within this NCTCT10442 type I SCCmec element is believed to produce a non-functioning recombinase [22, 23]. Given that all the isolates in this collection have this frameshift mutation, this strongly supports the conclusions of the phylogenetic analysis, namely that a type I SCCmec was acquired once in the ST250 background, and then became fixed in the population due to defective recombinase apparatus that precluded excision.

One of the questions we sought to address in this study was what were the temporal events surrounding the emergence of MRSA. With the first reports of MRSA occurring only after introduction of methicillin in the UK in 1959 and Denmark in 1964 it seemed reasonable to conclude that resistance arose after the first clinical use of the drug, and resistance therefore developed in *S. aureus* as an adaptive response following exposure to the antibiotic. However, the results presented in this communication are not consistent with this conclusion, since the gene bestowing methicillin

resistance was likely to have been acquired in the mid-1940s. It was during this period that  $\beta$ -lactamase-mediated penicillin resistance was becoming widespread among clinical isolates of *S. aureus*. Within 4 years of the introduction of penicillin for the treatment of staphylococcal infections, the first penicillin-resistant *S. aureus* were being described in 1944 [1]. In the years that followed the frequency of resistance in clinical isolates climbed steadily, such that by the time methicillin was introduced into clinical practice in 1960, resistance rates of 80% were common [24, 25].

Whilst the main genetic determinant associated with penicillin resistance in *S. aureus* is *blaZ*, *mecA* also encodes penicillin resistance via a different mechanism involving an alternative penicillin-binding protein, PBP2a [6, 26]. In the sequenced collection *blaZ* is widely distributed, albeit at a lower frequency than *mecA* (85.2% of isolates carry the *blaZ* gene in comparison to 95.2% for *mecA*), suggesting a selective advantage to possessing two distinct  $\beta$ -lactam resistance mechanisms. Based on the temporal calibration of the acquisition of *mecA*, it appears likely that methicillin resistance in *S. aureus* evolved long before this new  $\beta$ -lactam antibiotic was introduced. Thus, it was the widespread use of penicillin, rather than methicillin, that was the driver for the emergence in the archaic MRSA clone.

Beyond  $\beta$ -lactams our analysis uncovered evidence for the strong selective impact that a number of different antibiotics have had on the evolution of the archaic

MRSA clone. Several of the antibiotics, such as tetracycline, are prescribed in far lower amounts today in human medicine than in the 1950s and 1960s, and resistance to these antibiotics in contemporary *S. aureus* from humans is relatively rare, which contrasts with the archaic MRSA population, in which the distribution of tetracycline resistance determinants was widespread (Fig. 2; 96% of isolates contained *tetK* or/and *tetM*) [27]. In a prescient study examining the antibiotic consumption and rates of resistance in a hospital in the US in the 1950s, Bauer et al. provided evidence of a correlation between the two, where increasing usage of tetracycline was associated with an increase in the frequency of tetracycline resistance in isolates from inpatients [25].

In addition to methicillin and tetracycline resistances, a key phenotypic marker of the archaic MRSA clone was non-susceptibility to streptomycin. In our analysis we identified a mutation predicted to confer streptomycin resistance occurring on the same branch of the tree in which we mapped the acquisition of the *SCCmec* element. This finding suggests that methicillin and streptomycin resistance both emerged in the archetypal MRSA progenitor population around the same time. Discovered in the early 1940s, streptomycin was demonstrated to have activity against Gram-positive pathogens, and was used in the UK in 1947 during the first ever randomized clinical trials studying the efficacy of streptomycin in the treatment of pulmonary tuberculosis [28, 29]. It therefore appears that the first MRSA clone emerged, and developed resistance to two of the earliest antibiotics—streptomycin and penicillin—almost immediately after the *S. aureus* population would have been first exposed to them.

At the time of its discovery, the incidence of MRSA in the general population is likely to have been very low. This is demonstrated by the fact that screening of over 5000 samples at Public Health England yielded only three methicillin-resistant isolates. Therefore, it is likely that when methicillin was introduced to circumvent penicillin resistance in *S. aureus*, it did not select for emergence of MRSA at that time, but instead provided the selective pressure, which drove the nosocomial spread of a pre-existing variant, at a time when infection control measures in UK hospitals were limited.

## Conclusions

This study highlights the unintended consequences of widespread antibiotic use, and how when new drugs are introduced to bypass known resistance mechanisms, they may already be rendered ineffective due to unrecognised adaptations accrued in response to prior selective pressures exerted by other antibiotics. This remains one of the many challenges in tackling the growing problem of AMR and serves to emphasise the importance of continual surveillance of pathogen populations for evidence

of emerging adaptations and resistance patterns in the context of prescribing practice.

## Methods

### Bacterial isolates

Two hundred and nine isolates derived from the culture collections of *Staphylococcus aureus* reference laboratory, Public Health England, and isolates originating from the Statens Serum Institute collected and analysed by Profs Tomasz, Westh and de Lencastre. These correspond to a collection of MRSA and MSSA isolates collected between 1960 and the late 1980s in the UK and Denmark. Isolates from the Public Health England collection were all retrieved from the original freeze-dried cultures put down in the 1960s. All the Statens Serum Institute isolates were kept in a lyophilized state until the late 1990s, when they were opened, sub-cultured, and then stored at  $-80^{\circ}\text{C}$ . A record of the number of sub-cultures prior to freeze-drying, or post-isolate recovery, was not available, but it is known that the isolates were not repeatedly sub-cultured over the last 50 years.

One hundred and eighty eight isolates preserved as freeze-dried cultures in the Health Protection England (HPE) Staphylococcal Reference Laboratory were resurrected and grown on solid media. Prior to the start of this study the Reference Laboratory sequence typed all isolates from 1960 and 1961 using standard MLST techniques [30] and identified that the isolates belonged to CC8 and were either ST250 or ST247.

Twenty-one CC8 MRSA isolated in Denmark between 1964 and 1989 were also included in this study. These isolates originating from the Statens Serum Institute and had been previously sequence typed using standard MLST techniques [30]. All isolates in this study were subsequently sequence typed from their whole genome sequence data (see below).

### Genomic library preparation and sequencing

Genomic DNA was isolated using the Qiagen QIAcube system according to the manufacturer's protocol.

We prepared sequencing libraries from 500 ng of DNA extracted from each MRSA isolate as previously described, with amplification using Kapa HiFi polymerase (Kapa Biosystems, Woburn, MA, USA) [31]. Tagged DNA libraries were created using a method adapted from a standard Illumina Indexing protocol, as described previously [31]. Whole genome sequencing was performed on the Illumina HiSeq 2000 platform with 100-bp paired-end reads. The Illumina sequence data have been submitted to the European Nucleotide Archive (ENA) and the accession numbers are provided in Additional file 1: Table S1.



### Bioinformatic and phylogenetic analysis

The sequence reads for each representative isolate ( $n = 209$ ) were aligned against the reference genome of the MRSA *S. aureus* COL (accession number CP000046) [32] using SMALT (version 0.7.4; <http://www.sanger.ac.uk/science/tools/smalt-0>) and SNPs (single nucleotide polymorphisms) and indels (insertions/deletions) identified as described previously [31]. Mobile genetic elements (MGEs) were identified in the COL reference chromosome by comparison with other *S. aureus* chromosomes, where BLASTN (version 1.4) pairwise comparison were visualized in ACT (version 13.0.0) [33]. Regions of recombination within core genome and SCCmec element alignments were identified with Gubbins using the default parameters (version 1.4.10; <http://github.com/sanger-pathogens/Gubbins>) [34]. Phylogenetic reconstruction using core SNPs was performed with RAxML (version 8.2.8), using a GTR model with a gamma correction for among site rate variation [35]. Regions of high-SNP density corresponding to putative regions of recombination and those SNPs associated with horizontal gene transfer were excluded. Assembly of all genomes was performed using a high throughput assembly method [36].

In order to investigate if the genomic data contained evidence of a temporal signal we used root to tip linear regression using Path-O-Gen (version 1.4; <http://tree.bio.ed.ac.uk/software/tempest/>; Additional file 3: Figure S4). A core alignment for 122 isolates for which precise dates of isolation were available was used. MGEs and regions of predicted recombination along with homoplasic SNPs within these isolates were then also excluded. To estimate evolutionary rates and time to most common recent ancestor (TMRCA) Bayesian phylogenetic reconstruction was performed using BEAST (version 1.7.4) [37]. A GTR model with a gamma correction for among-site rate variation was used, and all combinations of strict, relaxed lognormal, and relaxed exponential clock models and constant, exponential, expansion, and skyline population models were evaluated. For each, three independent chains were run for 100 million generations, sampling every ten generations. On completion each model was checked for convergence, both by checking effective sample size (ESS) values were greater than 200 for key parameters, and by checking independent runs had converged on similar results. Models were compared for their fit to the data using Bayes factors based on the harmonic mean estimator as calculated by the program Tracer (version 1.4) from the BEAST package. A burn-in of ten million states was removed from each of the three independent runs of this model before combining the results from those runs with the logcombiner program from the BEAST package.

A previously described database of sequences of known resistance determinant genes, both horizontally

acquired and core, was utilized as a resistome database (Additional file 4: Table S3) [27, 38]. Fastq files from the 209 isolates were mapped to the resistome database with SRST2 (version 0.1.8) using the default settings [39]. SNPs in chromosomally encoded genes previously identified as being associated with antimicrobial resistance were then manually inspected to confirm the variation.

The multilocus sequence type (MLST) of isolates was predicted using SRST2 (version 0.1.8) [39].

### Additional files

**Additional file 1: Table S1.** Isolate metadata. The table contains information about the origins, mapping and assembly statistics, and genotype information of all the isolates used in the study. For the origins of the isolates, the year or decade of isolation, along with place and country of isolation are provided where known. Isolates designated London originated from multiple hospitals (hospitals 1–5), and these are indicated. Isolates that were described in the original description by P. Jevons, published in the BMJ in 1961 are highlighted. Summary information for the *de novo* assembly and reference mapping of each isolate's whole genome sequence data is provided, as is the ENA accession number for fastq data. The sequence type (ST) and allele type are provided, and single locus variants (SLV) or a triple locus variant (TLV) are marked. The isolates used in the temporal analysis presented in Figs. 3b and 4 are marked. (XLSX 90 kb)

**Additional file 2: Table S2.** Isolates from the original description of MRSA. Minimum inhibitory concentration (MIC) to celbenin (methicillin) derived from the original description by P. Jevons, published in the *BMJ* in 1961. MIC values represent the variation noted between colonies. Expected range of sensitivity to celbenin in coagulase-positive staphylococci 1.25–2.5 µg/ml. Acquired antibiotic resistance genes and core resistance mutations identified in the genomes are indicated. (PDF 42 kb)

**Additional file 3: Figure S1.** Maximum likelihood tree of SCCmec type I elements in historic MRSA isolates. **Figure S2.** Posterior support of maximum clade credibility trees of the historic MRSA population based on BEAST analysis (as illustrated in Fig. 3b). **Figure S3.** Linear regression of the root-to-tip distances of historic MRSA SCCmec type I elements. **Figure S4.** Linear regression of the root-to-tip distances of the archetypal MRSA clone population used for BEAST analysis. (PDF 302 kb)

**Additional file 4: Table S3.** Resistome antibiotic determinants. Updated resistome database used for the *in silico* identification of antibiotic-resistant determinants based on the database previously described by Aanensen et al., 2016. Included in the database are antibiotics and associated acquired genes and core gene mutations. Accession numbers for the acquired genes and protein sequences are provided, as well as amino acid substitution information for the core genes. (PDF 80 kb)

### Abbreviations

AMR: Antimicrobial resistance; CC: Clonal complex; CoNS: Coagulase negative staphylococci; CWA: Cell wall-anchored; ENA: European Nucleotide Archive; HPD: Highest posterior density; MGE: Mobile genetic element; MIC: Minimum inhibitory concentration; MLST: Multi-locus sequence typing; MRSA: Methicillin-resistant *Staphylococcus aureus*; MSSA: Methicillin-sensitive *Staphylococcus aureus*; PBP: Penicillin binding protein; SCCmec: Staphylococcal cassette chromosome *mec*; SLV: Single locus variant; SNP: Single nucleotide polymorphism; ST: Sequence type; TMRCA: Time to most recent common ancestor

### Acknowledgements

We thank Nick Thomson for reading the manuscript and useful discussions. We also thank the core sequencing and informatics teams at the Sanger Institute for their assistance and The Wellcome Trust for its support of the Sanger Institute Pathogen Genomics and Biology groups.

## Funding

SDB, JP and MTGH were supported by Wellcome Trust grant 098051. CPH was supported by Wellcome Trust grant number 104241/z/14/z. Bioinformatics and computational biology analyses were supported by the University of St Andrews Bioinformatics Unit, which is funded by a Wellcome Trust ISSF award (grant 097831/Z/11/Z). AK, MD and BP received funding from Public Health England.

## Availability of data and materials

Short reads for all sequenced isolates have been submitted to the European Nucleotide Archive (ENA; <http://www.ebi.ac.uk/ena/>) under study accession number ERP001103.

## Authors' contributions

MTGH and CPH carried out data analyses, interpreted the data, and wrote the paper. BP and MD carried out data analyses and cultured the isolates and helped write the paper. SDB jointly conceived the project with AMK, and with AT and HdL. JP facilitated sequencing of the isolates, and helped write the paper. HW, AT, HdL, SDB, and AMK provided isolates and helped write the paper. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Ethics approval and consent to participate

No ethics approval was required for this work.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>School of Medicine, University of St Andrews, St Andrews KY16 9TF, UK.

<sup>2</sup>School of Medicine, University of Dundee, Dundee DD1 9SY, UK.

<sup>3</sup>Antimicrobial Resistance and Healthcare Associated Infections Reference Unit, National Infection Service, Public Health England, Colindale, UK. <sup>4</sup>The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK. <sup>5</sup>Department of Clinical Microbiology, Hvidovre University Hospital, Hvidovre, Denmark. <sup>6</sup>Institute of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark. <sup>7</sup>Laboratory of Microbiology and Infectious Diseases, The Rockefeller University, New York, USA. <sup>8</sup>Laboratory of Molecular Genetics, Instituto de Tecnologia Química e Biológica António Xavier (ITQB), Universidade Nova de Lisboa, Oeiras, Portugal.

Received: 9 April 2017 Accepted: 7 June 2017

Published online: 20 July 2017

## References

- Kirby WM. Extraction of a highly potent penicillin inactivator from penicillin resistant staphylococci. *Science*. 1944;99:452–3.
- Knox R. A new penicillin (BRL 1241) active against penicillin-resistant staphylococci. *Br Med J*. 1960;2:690–3.
- Jevons MP. "Celbenin"-resistant Staphylococci. *Br Med J*. 1961;1:124–5.
- Eriksen KR, Erichsen I. Resistance to methicillin, isoxazolyl penicillins, and cephalothin in *Staphylococcus aureus*. *Acta Pathol Microbiol Scand*. 1964;62: 255–75.
- Katayama Y, Ito T, Hiramatsu K. A new class of genetic element, staphylococcus cassette chromosome *mec*, encodes methicillin resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother*. 2000;44:1549–55.
- Hartman BJ, Tomasz A. Low-affinity penicillin-binding protein associated with beta-lactam resistance in *Staphylococcus aureus*. *J Bacteriol*. 1984;158: 513–6.
- Matthews P, Tomasz A. Insertional inactivation of the *mec* gene in a transposon mutant of a methicillin-resistant clinical isolate of *Staphylococcus aureus*. *Antimicrob Agents Chemother*. 1990;34:1777–9.
- Crisóstomo MJ, Westh H, Tomasz A, Chung M, Oliveira DC, de Lencastre H. The evolution of methicillin resistance in *Staphylococcus aureus*: similarity of genetic backgrounds in historically early methicillin-susceptible and -resistant isolates and contemporary epidemic clones. *Proc Natl Acad Sci U S A*. 2001;98:9865–70.
- Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H, Spratt BG. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). *Proc Natl Acad Sci U S A*. 2002;99:7687–92.
- Oliveira DC, Tomasz A, de Lencastre H. Secrets of success of a human pathogen: molecular evolution of pandemic clones of methicillin-resistant *Staphylococcus aureus*. *Lancet Infect Dis*. 2002;2:180–9.
- Gomes AR, Westh H, de Lencastre H. Origins and evolution of methicillin-resistant *Staphylococcus aureus* clonal lineages. *Antimicrob Agents Chemother*. 2006;50:3237–44.
- Dyke KGH, Jevons MP, Parker MT. Penicillinase production and intrinsic resistance to penicillins in *Staphylococcus aureus*. *Lancet*. 1966;287:835–8.
- Ito T, Katayama Y, Asada K, Mori N, Tsutsumimoto K, Tienasitorn C, et al. Structural comparison of three types of staphylococcal cassette chromosome *mec* integrated in the chromosome in methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother*. 2001;45:1323–36.
- Huang H-N, Han J-X, Wang J-H, Song C-Z, Zhang Z-L, Liang H. *rpsL* gene analysis associated with streptomycin resistance in *Mycobacterium tuberculosis*. *Yi Chuan Xue Bao*. 2003;30:376–81.
- Murphy E, Huwylar L, de Freire Bastos MDC. Transposon Tn554: complete nucleotide sequence and isolation of transposition-defective and antibiotic-sensitive mutants. *EMBO J*. 1985;4:3357–65.
- Couto I, de Lencastre H, Severina E, Kloos W, Webster JA, Hubner RJ, et al. Ubiquitous presence of a *mecA* homologue in natural isolates of *Staphylococcus sciuri*. *Microb Drug Resist*. 1996;2:377–91.
- Kreiswirth BN, Kornblum J, Arbeit RD, Eisner W, Maslow JN, McGeer A, et al. Evidence for a clonal origin of methicillin resistance in *Staphylococcus aureus*. *Science*. 1993;259:227–30.
- Shore A, Rossney AS, Keane CT, Enright MC, Coleman DC. Seven novel variants of the staphylococcal chromosomal cassette *mec* in methicillin-resistant *Staphylococcus aureus* isolates from Ireland. *Antimicrob Agents Chemother*. 2005;49:2070–83.
- Juuti KM, Sinha B, Werbeck C, Peters G, Kuusela PI. Reduced adherence and host cell invasion by methicillin-resistant *Staphylococcus aureus* expressing the surface protein PIs. *J Infect Dis*. 2004;189:1574–84.
- Santos-Júnior CD, Veríssimo A, Costa J. The recombination dynamics of *Staphylococcus aureus* inferred from *spa* gene. *BMC Microbiol*. 2016;16:143.
- Murphy E, Lin SL, Nunez L, Andrew L, Fink PS, Dilts DA, et al. Challenges for the evaluation of *Staphylococcus aureus* protein based vaccines: monitoring antigenic diversity. *Hum Vaccin*. 2011;7(Suppl):51–9.
- Misiura A, Pigli YZ, Boyle-Vavra S, Daum RS, Boocock MR, Rice PA. Roles of two large serine recombinases in mobilizing the methicillin-resistance cassette *SCCmec*. *Mol Microbiol*. 2013;88:1218–29.
- Noto MJ, Archer GL. A subset of *Staphylococcus aureus* strains harboring staphylococcal cassette chromosome *mec* (SCCmec) type IV is deficient in CcrAB-mediated SCCmec excision. *Antimicrob Agents Chemother*. 2006;50: 2782–8.
- Barber M, Rozwadowska-Dowzenko M. Infection by penicillin-resistant staphylococci. *Lancet*. 1948;2:641–4.
- Bauer AW, Perry DM, Kirby WM. Drug usage and antibiotic susceptibility of staphylococci. *JAMA*. 1960;173:475–80.
- Ubukata K, Nonoguchi R, Matsushashi M, Konno M. Expression and inducibility in *Staphylococcus aureus* of the *mecA* gene, which encodes a methicillin-resistant *S. aureus*-specific penicillin-binding protein. *J Bacteriol*. 1989;171:2882–5.
- Aanensen DM, Feil EJ, Holden MTG, Dordel J, Yeats CA, Fedosejev A, et al. Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive *Staphylococcus aureus* in Europe. *MBio*. 2016;7:e00444–16.
- Schatz A, Bugie E, Waksman SA. Streptomycin, a substance exhibiting antibiotic activity against Gram-positive and Gram-negative bacteria. *Proc Soc Exp Biol Med*. 1944;55:66–9.
- Anonymous. Streptomycin treatment of pulmonary tuberculosis. *Br Med J*. 1948;2:769–82.
- Enright MC, Day NPJ, Davies CE, Peacock SJ, Spratt BG. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J Clin Microbiol*. 2000;38:1008–15.
- Hsu L-Y, Harris SR, Chlebowicz MA, Lindsay JA, Koh TH, Krishnan P, et al. Evolutionary dynamics of methicillin-resistant *Staphylococcus aureus* within a healthcare system. *Genome Biol*. 2015;16:81.
- Gill SR, Fouts DE, Archer GL, Mongodin EF, Deboy RT, Ravel J, et al. Insights on evolution of virulence and resistance from the complete genome

- analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. *J Bacteriol.* 2005;187:2426–38. <http://jb.asm.org/cgi/doi/10.1128/JB.187.7.2426-2438.2005>.
33. Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the Artemis Comparison Tool. *Bioinformatics.* 2005;21:3422–3.
  34. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 2014;43:e15.
  35. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30:1312–3.
  36. Page AJ, De Silva N, Hunt M, Quail MA, Parkhill J, Harris SR, Otto TD, Keane JA. Robust high-throughput prokaryote de novo assembly and improvement pipeline for Illumina data. *Microbial Genomics.* 2016;2:e000083.
  37. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol.* 2012;29:1969–73.
  38. Holden MTG, Hsu L-Y, Kurt K, Weinert LA, Mather AE, Harris SR, et al. A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant *Staphylococcus aureus* pandemic. *Genome Res.* 2013;23:653–64.
  39. Inouye M, Dashnow H, Raven L-A, Schultz MB, Pope BJ, Tomita T, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med.* 2014;6:90.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

